The Intergenerational Transmission of Adiposity across Countries

Peter Dolton  
(University of Sussex)  
Centre for Economic Performance  
London School of Economics)  

and  

Mimi Xiao  
(University of Sussex)  

Abstract  
There is a worldwide epidemic of obesity. We are only slowly beginning to understand its causes and consequences. This paper addresses one important component of the crisis – namely the extent to which obesity – or more generally - adiposity is passed down from one generation to the next. Using the BMI as a measure of adiposity, we find that the intergenerational transmission of adiposity is relatively high and very comparable across time and countries – even if these countries are at a very different stage of economic development. Our second key finding is that this intergenerational transmission mechanism is very different across the distribution of children’s BMI. Most specifically it is up to double for the fattest children what it is for the thinnest children. This has enormous consequences for the health of the world’s children.

Address for Correspondence:  
Prof Peter Dolton  
Department of Economic  
University of Sussex,  
Brighton  
BN1 9SL
**Introduction**

There is a worldwide epidemic of obesity and obesity has become one of the foremost major public health problem in most countries. In 2013, the US spent 190 billion dollars on obesity-related health expenses. We are only slowly beginning to understand the causes and consequences of childhood obesity. This paper addresses one important component of the crisis – namely the extent to which obesity – or more generally - adiposity is passed down from one generation to the next.

Hence, our underlying question is: what is the driving force behind rising childhood obesity? Adiposity is a result of both the genetic process of inheritance and a result of decisions made in families. To understand the process of obesity it is crucial to understand the intergenerational transmission mechanism behind this. Evidence (Maes, Neale, & Eaves, 1997) suggests that adiposity is affected by both environmental and genetic factors. Clearly, the intergenerational transmission mechanism we are studying here operates through both these two channels. So it is transmitted through family environmental factors, which take the form of intra-household mechanism (how the resources are allocated within the family), and it is also affected by genetic factors through a direct channel. Therefore, through exploring the elasticity of adiposity (measured by the BMI score) across generations in different countries, we attempt to reveal the underlying intergenerational relationship in anthropometric characteristics.

In order to give you a sense of what the underlying relationship between parents and child’s BMI looks like – we first of all present some basic non-parametric graphs of the data. First we plot a non-parametric scatterplot based on the raw data. Figure 1 below is the local weighted scatter smoothening of the log of father’ BMI variable against the log of their children’s BMI variable, the slopes capture the magnitude of the intergenerational elasticity. They suggest that the slopes are relatively flat and nearly parallel across countries, this implies that the intergenerational elasticity are relatively high and approximately constant across countries.
Figure 1a Lowess log (father’s BMI variable) and log (child’s BMI variable)

Figure 1b Lowess log (mother’s BMI variable) and log (child’s BMI variable)
2. Evidence on the Intergenerational Transmission Mechanism.

Intergenerational studies originate with Francis Galton (1877). By running a regression of the offspring’s height on their parents’ height, he argued that an individual’s characteristics are correlated with those of their parents and at the same time “regress to mediocrity”. More specifically, the individual characteristics (such as height and weight) are closer to the population mean than those of their parents (Galton, 1892). This finding was the basis of Becker-Tomas model (1979, 1986) on intergenerational human capital transmission (Goldberger, 1989; Han & Mulligan; Mulligan, 1999).

Most of these intergenerational transmission studies are about the transmission of income or educational achievement outcomes. Since this transmission relates to the equality (or otherwise) of individual opportunity over time, which exerts profound influences on the social mobility. The strength of this transmission is usually measured by the elasticity of children’s income with respect to their parents’ income (or the intergenerational elasticity of income, hereafter called IIE). The larger is the IIE the more it means that the children’s position on the “income ladder” is determined by their parents’ income position. We would naturally be concerned if this elasticity of the transmission mechanism was (too) large as it would imply that both equity and efficiency of the society would be undermined. Specifically, it means that the smaller is this elasticity the smaller the role played by one’s parents in the determination of the child’s outcome in school or the labour market.

In terms of the discrepancy in IIE across countries, partly due to the restriction of data which covers multiple generations, most of these studies are conducted in the US or European countries. In the US, the consensus estimated IIE is “0.4 or a bit higher” (for instance, 0.473 using PSID by Grawe (2004), 0.542 using NLSY sample born 1957-64 by Bratsberg et al. (2007), this is higher than Canada (0.2 using register data (Corak et al., 1999), 0.152 using IID Canadian Intergenerational Income Data and 0.381 using PSID Panel Study of Income Dynamics data (Grawe, 2004) and most of the European countries except for Britain (0.45 using NCDS 1958 cohort Bratsberg, et al., 2007) and Italy (0.48 using data Italy Survey on Household Income and Wealth (SHIW), Piraino, 2007). The IIE estimates in Nordic countries or Scandinavian societies may be the lowest, ranging from 0.2 to 0.3 (see Pekkarinen, et al., 2009; Björklund, et al., 2012). In contrast, the IIE in China is perhaps at the top of the list---0.63: a Chinese father’s income 10 percent above the paternal cohort
mean will be associated with his son having an income 6.3 percent above the filial cohort mean (Gong, 2012). Using the Urban Household Education and Employment Survey (UHEES) and the Urban Household Income and Expenditure Survey in 1987-2004 (UHIFS), Gong et al. (2012) show the IIE in China is 0.63 for father-son, 0.97 for father-daughter, 0.36 for mother-son, and 0.64 for mother-daughter, and education is one of the most crucial channels through which earnings ability is transmitted across generations. However, other factors such as genes and health are also potentially important pathways of intergenerational transmission in China.

Regarding the intergenerational transmission of education achievement, the intergenerational mechanism can be written in the following formula, in which case the child’s education achievement is measured by the human capital: \( H_c = F(Y_c, H_p, A_c) \), thus the mechanism usually uses three main channels. The first channel is through the effects of parental income, higher educated parents tend to have more income, so they have more resources to invest in child’s education \( Y_c \); the second is through the effects of parental education \( H_p \), since higher-educated parents may invest in child’s education in a more efficient way. In addition to these two indirect channels, parental education may also affect child’s education through a direct channel, which is usually proxied by the genetic inheritance of ability \( A_c \). Empirically, the first channel can be decomposed into the effects of current parental income and the effects of permanent parental income, of which the latter normally plays the dominant role and might be measured by family fixed effects (Carneiro and Heckman, 2003). The second channel also includes the motivation effects since higher educated parents may encourage their children to go for a higher level of education (Boudon, 1974). The third channel is normally conducted by comparing children of twin pairs (Behrman and Rosenzweig, 2002) or between biological and adopted children with variation in education (Björklund et al., 2006), the general conclusion is that the intergenerational correlation in education cannot be fully attributed to the genetic factors. The intergenerational correlations of education is estimated at 0.3 ~ 0.4 (Alburg, 1998), but a more popular measure of this intergenerational educational relationship is intergenerational education elasticity (hereafter called IEE), which varies from 0.14~ 0.45 in the USA (Mulligan, 1999) to 0.25 ~0.4 in the UK (Dearden et al., 1997). It is worth mentioning that a large body of the literature looks at the intergenerational elasticity of IQ (hereafter called IQE), which is used to measure the
intergenerational relationship in the third channel, the magnitude of IQE ranges from 0.3 to 0.5 (Solon, 2004; Anger and Heineck, 2010; van Leeuwen et al., 2008).

There is also a growing literature on the intergenerational correlation in various health outcomes, such as weight, height, BMI, self-rated measures (Coneus and Spieß, 2008), depression (Akbulut and Kugler, 2007), and smoking behaviours (Loureiro et al., 2006). Using data from the German Socio-Economic Panel (SOEP), Coneus and Spieß (2008) estimate the intergenerational relationship of both father and mother and children. Regarding the fixed effects estimates, using actual BMI and obesity as anthropometric measures, they find that father’s BMI has a significantly positive effect on child’s BMI (with a coefficient of 0.57, but the effects of mother’s is not significant), whereas mother’s obesity is strongly associated with child’s obesity with a coefficient of 0.26. They claim this as a “transmission” rather than merely “link”. However, in their data, child’s health outcomes are provided by mother rather than professionals, additionally, father and mother’s health are self-reported, this might lead to biased estimates from measurement error. Classen (2010) estimates the intergenerational elasticity of BMI when both generations are between the ages of 16 and 24. Applying a regression which only controls for mother and child’s BMI, he finds an elasticity of 0.35 between mother and child’s BMI, but typically this sort of long data is not available. As Black and Devereux (2010) review, few papers have claimed a causal link, since the family environmental factors may affect the health outcome of both parents and children. Some studies try to address this by differencing out fixed family characteristics through comparing “sibling mothers” or looking within “twin pairs of mothers”, assuming twins share the same environment and genetics (Currie and Moretti, 2007; Black, Devereux, and Salvanes, 2007; Royer, 2009).

In the case of income and education, the intergenerational elasticity varies largely across countries, and the environmental channels are exploited more often than the genetic channel. Whether this transmission mechanism is applicable to the case of anthropometric outcome (measured by BMI variable in this paper) generates the main motivation of our study. The difference of IBE from IIE and IEE hinges on the relative role and the interaction of environmental and genetic forces in the intergenerational transmission, our assumption is that in the transmission of BMI variable, a smaller fraction of the operation forces are open to manipulation (such as the diet change with the household--environmental change), and a
larger fraction of the forces are driven by the “natural process”. In other words, in the case of health outcome such as the BMI variable, they are more likely to be inherited genetically regardless of the change in environment. If this assumption is true, our estimation for IBE may provide a lower bound of the intergenerational correlation in any characteristics including the income and education. In other words, assuming the intergenerational transmission of anthropometric outcome is entirely determined by the genetic traits, if our IBE is closer to the IIE in Scandinavian societies where the IIE is the lowest---0.2, it may imply that the relationship between parents and child cannot be lower than this threshold, in spite of the change in either family environment (such as the shift of nutrition pattern) or socioeconomic environment (such as the innovation or marketing campaign in food industry).

In addition to “regression to the mean” in the inheritability of BMI variable, the degree of this inheritability (IBE) may vary across child’s BMI distribution and this variation usually relates to the family’s socioeconomic status in the society. The general conclusion in the literature is: in either developed countries or developing countries, the intergenerational correlation in health measure tends to be stronger at lower SES levels (see, for example, Currie and Moretti, 2007; Bhalotra and Rawlings, 2009). In developing countries, this strong correlation emerges at the lower levels of BMI, whereas in developed countries such as the US, this also occurs at higher levels of BMI (Classen, 2010, Laitinen et al., 2001; Scholder et al., 2012), one explanation is that in these countries, fast food industry is more developed, these “unhealthy” food are generally cheaper than “healthy” food, thus lower income families tend to consume these “unhealthy” food which is viewed as a cause of obesity.

Thus far, we can see these intergenerational studies are essentially derived from the model of “Regression to the mean” (Becker or even dates back to Galton). Based on this framework, the intergenerational transmission of BMI variable in this paper can be modelled in the following way:

$$\log(BMI_c) = \sigma + h\log(BMI_p) + v_c$$  \hspace{1cm} (1)

Where $BMI_c$ denotes child’s BMI variable, $BMI_p$ represents parents’ BMI variable (father’s or mother’s or both), and $v_c$ denotes the random determinants of $BMI_c$. Thus,
$h$, the intergenerational elasticity of BMI variable (hereafter called IBE) measures the degree of inheritability of BMI variable if we assume BMI is determined genetically as the innate ability in Becker’s innate ability model.

### 3. An Empirical Model of Intergenerational BMI Transmission

In this section, we outline an empirical model on intergenerational transmission of BMI. This model is directly analogous to Becker’s model on intergenerational transmission of income. In Becker’s model, parents allocate their income between the child’s income, and his own consumption, to optimize their utility. In our model, the outcome of interest is a child’s BMI health which can be invested by parents sacrificing their own consumption (literally—maybe even their own food)—to promote the improved health—as measured by BMI of the child. Hence here, $Y$ denotes the child’s health as the intergenerational outcome we are interested in. The child’s health is a function of parent’s income, $X$, and a genetic endowment, $E$, which is determined exogenously. Since children cannot choose their parents and the genetic traits they inherit from them, then this endowment factor is taken as given (by the exogenous “natural selection process”).

\[ Y = bX + aE \]  
\[
\text{(1)}
\]

Where $E$ is decomposed into genetic factors, $e$, and environmental factors, $u$.

\[ E = e + u \]  
\[
\text{(2)}
\]

From this point on, we will use lower case letters to denote observable variables which we obtain data on or can proxy for. Let subscript $p$ index the parent and $i$ index the child, substituting (2) into equation (1), we obtain

\[ y_i = \alpha x_p + b e_i + u_i \]  
\[
\text{(3)}
\]
As for the correlation between parents’ genes and the child’s genes, we consider the child’s genes, $e_i$, as a production function of its father’s genes, $e_{fi}$, and its mother’s genes, $e_{mi}$, which is assumed to take the form of Cobb-Douglas function\(^1\).

$$e_i = Ae_{fi}^\alpha e_{mi}^\beta v_p$$

(4)

Where $A$ denotes the child’s multiplicative scaled genetic transmission and $v_p$ denotes the stochastic term. Equation (4) can be thought of as the ‘translog biological production function’, which maps parents’ health outcomes into children’s outcomes.

$$\log(e_i) = \log(A) + \alpha \log(e_{fi}) + \beta \log(e_{mi}) + \log(v_p)$$

(5)

Now consider substituting equation (5) into equation (3) where we assume that mothers and fathers’ BMI measures are sufficient statistics for their health and the environmental factors are individual specific and captured by the term $f_i$. We now wish to estimate the following equation (6) in a cross-section framework.\(^7\)

$$\log(y_i) = \delta + \alpha \log(y_{fi}) + \beta \log(y_{mi}) + \gamma x_p + f_i + \varepsilon_i$$

(6)

where $i$ indexes individual child observations and $\varepsilon_i$ captures the transmitted stochastic error term. Equation (6) shows that child’s health outcome $y_i$ is a function of child $i$’s father’s health outcome, $y_{fi}$, and the child’s mother’s health outcome, $y_{mi}$, $x_p$ denotes a vector of parental income factors (or family environmental factors), such as the categories of household income per capita, father and mother’s SES factors. As both child’s health and parents’ health are affected by time-invariant unobserved individual heterogeneity, $f_i$, such as eating habits, health behavior and genetic components. we estimate equation (6) in an individual fixed effect framework.

$$\log(y_{it}) = \delta + \alpha \log(y_{fit}) + \beta \log(y_{mit}) + \gamma x_{pt} + f_i + \varepsilon_{it}$$

(7)

\(^1\)The justification of this functional form is purely for analytical tractability, and is not strict necessary. One similar specification is in Rosenzweig, (1983), where birth weight is a function of health behavior such as smoking behaviour of the parent.
Where \( t \) denote observations referenced to a specific time period (or wave of the data). This equation takes into account an individual fixed effect \( f_i \). Based on this equation and using BMI as the measure of health, we estimate the intergenerational elasticity of health outcome, conditional on time-variant family environmental factors which are captured by \( x_p \).

However, it is also reasonable to assume that unobserved family environmental factors remain constant. For instance, the family routine, such as eating and sleeping time shared among household members, normally do not change over time; More importantly, the pattern of food allocation among household members normally remain constant, these patterns, together with, who is in control of the family income (Thomas, 1990), who takes a higher level of energy-intensive activities (Pitt et al., 1990), whether the parents have a preference for sons (Qian, 2008) or lower-birth-order children (Dasgupta, 1995), can affect both parental and children’s BMI outcome. Thus, household fixed effects are applied to estimate equation (6), ie, fixed effects model is estimated using the following equation (8).

\[
\log(y_{ijt}) = \delta + \alpha \log(y_{fjt}) + \beta \log(y_{mjit}) + \gamma x_{pjt} + \delta_i + \epsilon_{it}
\]

Where \( j \) indexes household observations. In household \( j \), child \( i \)'s health is a function of father’s health, \( y_{fjt} \), and mother’s health, \( y_{mjit} \), \( \delta_i \) denotes the household fixed effects. This equation can only be identified when we have data on siblings for which the \( f_i \) effects are distinct and the \( \delta_i \) are the same. We have a subset of our data for which we can estimate this model- namely when we have data as more than one child in each household.

In the regression tables which follow we estimate single parent version of equation (7) and (8) by variously restricting \( \gamma = \alpha = 0 \), and \( \gamma = \beta = 0 \), then examining just \( \gamma = 0 \) to examine the pure effect of both parents but no conditioning factors (i simplified ‘ both parents equation’).

In order to investigate the simplistic dynamic pattern of child’s BMI measure, equation (9) is estimated by including the previous value of child’s BMI measure, \( y_{it-1} \). Here we can net out the individual unobserved heterogeneity when \( v_{it} = v_{it-1} \). Next, the age of the child is included, in the regression equation (10), to allow for the possible of bias associated with the use of the WHO software for China where children have comparatively lower than average
BMI, if this varies by age. The inclusion of the age of the child as a regressor also acts as a proxy for the length of time the child has been exposed to the treatment of the family environmental effect.

\[
\log(y_{it}) = \delta + \alpha \log(y_{ft}) + \beta \log(y_{mt}) + \gamma \log(y_{i,t-1}) + \epsilon_{it} \tag{9}
\]

\[
\log(y_{it}) = \delta + \alpha \log(y_{ft}) + \beta \log(y_{mt}) + \gamma \log(y_{i,t-1}) + \text{child's age} + \epsilon_{it} \tag{10}
\]

The empirical estimation will be conducted in several stages. First, we estimate the IBE at the central level. Single parent version of equation (7) and (8) (father-child IBE and mother-child IBE) are estimated using all the individual-wave observations, \(\log(y_{i,t-1})\) is also accounted for to investigate the dynamic change of child’s BMI variable, then equation (7), (8) are estimated in a fixed effects framework. Second, we estimate single parent version of equation (7) and (8) in terms of different levels of family social economic status, which includes different levels of family income, mother’s education, father’s occupation and the time duration when the family was in poverty, respectively. Third, applying single parent version of equation (7) and (8), we estimate the IBE across different quantiles of child’s BMI variable, and compare the results with those using simplified both parents equation and equation (9). Fourth, applying single parent version of equation (7) and (8), we estimate the IBE by age group. Finally, we repeat our estimation of the IBE on samples aged 16~18 years old to examine the probability that this relationship is structurally different when the children have become adults.


The most widely used measure of body fat, or adiposity, is the Body Mass Index (BMI) which is calculated using the following formula BMI = \(\frac{\text{weight(kg)}}{\text{height}^2(\text{cm})}\) * 10,000, since we have problems that when dealing with the measurement of Body Mass of children of different age, we will not use the BMI per se, but the z-score of the BMI as adjusted by age and gender of the child. This conversion can be made using the 2006 WHO Growth Standards for preschool
children and the 2007 WHO Growth Reference for school age children and adolescents. In Stata, this conversion is implemented using a program from the WHO website\(^2\).

As mentioned in the literature review, the majority of intergenerational studies use elasticity (eg. IIE and IEE) as a measure of the intergenerational relationship. To facilitate the comparison of our results on anthropometric with other intergenerational results, we also adopt elasticity as the measure and thus transform BMI z-score (ranging from -6 to 6) to BMI variable (which ranges from 1th to 100th) to obtain the logarithm form in the calculation of elasticity. In this way, this percentile indicates the relative position of individual BMI z-score within the CHNS sample, since BMI z-score is based on the comparison with the WHO population (the z-score normally distributed population constructed by the international reference population using ‘Anthro’), the BMI z-score percentile (hereafter called BMI variable) is a relative anthropometric indicator based on the comparison with the WHO population and then within the CHNS. Simply speaking, since the BMI z-score ranges from -6 to 6, and we can’t take the log of a negative number, we transform this into a percentile of BMI z-score to facilitate the taking of logs and the computation of an intergenerational BMI elasticity.

An important problem we face is exactly how we correlate a child’s BMI on their parents BMI. Clearly a child’s BMI is a function of their age and gender – so a simply correlation of child’s BMI against parents BMI would not allow for this factor. One way to examine the intergenerational transmission is to wait until the child is an adult and then correlate the two BMIs. This is what Classen did. There are two problems with this – firstly the logistic one is that there is very little data when one has the child’s height and weight observed when they are an adult – as well as having their parents height and weight at the same time. The other big problem with this is that we are mainly concerned with childhood obesity and so waiting until they are adult does not help us.

One way around these problems is to take the child’s BMI when they are young – use the WHO’s program to compute the child’s BMI z score which explicitly allows for both the

child’s age and gender. Once we have this z score we can then ask what would their BMI be with such a z score when they are adult. The assumption that we have to make here is that we are assuming that the child would remain in the same position in the distribution when they are adult as when they are a child. Note that by doing this we are not, de facto, assuming that this is what will happen to that child when they are an adult – but rather simply getting an estimate of what the adult BMI is consistent with a given z score for the child. Although this is a strong assumption there is only one other way to proceed.

This would be to simply use the child’s BMI (as it is- even if they are very young) as the dependent variable in a regression on the parents BMI on the assumption that if we control for the child’s age, gender, age squared and an interaction of the child with that of their gender then we would have conditioned out for the non-linear effect of age on gender\(^3\). We use both these methods as a robustness check on our findings. Fortunately they do not differ much in their findings – with the former method giving lower variance in the tails than the latter variable. We will therefore use the first method in each of our country datasets. We report the second method in an Appendix available on request for those interested.

In the course on doing this research we had considered if there was any alternative way of retrieving the IBE. We contemplated using the WHO to generate z scores or percentiles and using these logged metrics. It turns out that the estimation of the BMI elasticity is sensitive to any possible transformation of its scale. – ie. to z scores or percentiles. So keeping the analysis simple has many virtues. It turns out that estimating the model in the log of BMI or the BMI itself does not make much difference – the elasticity is slightly smaller when estimated without logging. But since taking logs allows albeit crudely – for general non-linearity in the data and has the nice property that is preserves the constant elasticity across the values of the BMI then we adopt it here. This means also that it forces the elasticity to be a constant – which has the virtue that its first derivative (and hence the elasticity) is constant across the whole range of the BMI score.

\(^3\) The weakness of this method is that we have to assume that we can net out for the whole non-linear process of the child’s BMI rising as they age.
5 Empirical Evidence of intergenerational transmission

5.1 OLS

Applying equation (1), we estimate the IBE on CHNS data, Indonesian Family Life Survey (IFLS) data, British 1970 cohort data, Health Survey for England (HSE) data and National Health and Nutrition Examination Survey (NHNAES), respectively.

Table 1a reports the results on IBE when equation (1) controls for father’s BMI variable alone. It suggests that the father-child IBE estimates range from 0.198 in Indonesian sample to 0.235 in BCS sample, and they do not vary substantially across countries (this is different from the previous studies on IIE and IEE). For the UK, the IBE estimate on BCS sample (0.235) is close to that from HSE sample (0.202) \(^4\). These results suggest that the responsiveness of child’s BMI variable to parents’ BMI variable is around 0.20 and the extent of this “inheritability” is quite constant across countries. In other words, if the father’s BMI variable is 50% above the mean of their generation, on average his child’s BMI variable would be around 10% (50% * 0.20) above the mean of the children’s generation, and this seems to be regardless of the ethnicity and economic growth level. In a similar way, Table 1b presents mother-child IBE estimates from these samples, and we see a similar pattern as given by father-child IBE estimates. In addition, comparing Table 1a and 1b, we can see that in general, the mother-child IBE estimates are greater than father-child IBEs, the only exception is found in CHNS sample. Next, we incorporate both father and mother’s BMI variables (\(\log(BMI_{f})\) and \(\log(BMI_{m})\)) into equation (1), and the results are reported in Table 1c. As we expect, once we control for both father and mother’s BMI variables, the sizes of paternal and maternal BMI effects shrink significantly compared with Table 1a and Table 1b, with the dominance of father’s BMI effects.

\(^4\) Notice the HSE was collected from 1995 to 2010, and the BCS 1970 survey tracks the cohorts born in 1970 until they reached 26 years (1996). This seems to tell us that the IBE in Britain was declining from 1970 to 2010, if HSE and BCS 1970 can talk to each other (by which I mean they have similar design and structure).
Table 1a father-child

<table>
<thead>
<tr>
<th></th>
<th>China</th>
<th>Indonesia</th>
<th>BCS</th>
<th>HSE</th>
<th>NHANES3</th>
</tr>
</thead>
<tbody>
<tr>
<td>logbmi_father</td>
<td>0.225***</td>
<td>0.198***</td>
<td>0.235***</td>
<td>0.202***</td>
<td>0.204***</td>
</tr>
<tr>
<td></td>
<td>(0.0154)</td>
<td>(0.0106)</td>
<td>(0.0104)</td>
<td>(0.00948)</td>
<td>(0.0171)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.356***</td>
<td>2.407***</td>
<td>2.348***</td>
<td>2.547***</td>
<td>2.603***</td>
</tr>
<tr>
<td></td>
<td>(0.0476)</td>
<td>(0.0324)</td>
<td>(0.0331)</td>
<td>(0.0312)</td>
<td>(0.0557)</td>
</tr>
<tr>
<td>Observations</td>
<td>13,990</td>
<td>18,570</td>
<td>21,505</td>
<td>26,316</td>
<td>6,515</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.024</td>
<td>0.024</td>
<td>0.035</td>
<td>0.022</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1b mother-child

<table>
<thead>
<tr>
<th></th>
<th>China</th>
<th>Indonesia</th>
<th>BCS</th>
<th>HSE</th>
<th>NHANES3</th>
</tr>
</thead>
<tbody>
<tr>
<td>logbmi_mother</td>
<td>0.186***</td>
<td>0.165***</td>
<td>0.201***</td>
<td>0.188***</td>
<td>0.166***</td>
</tr>
<tr>
<td></td>
<td>(0.0144)</td>
<td>(0.00877)</td>
<td>(0.00832)</td>
<td>(0.00789)</td>
<td>(0.0130)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.477***</td>
<td>2.500***</td>
<td>2.468***</td>
<td>2.600***</td>
<td>2.733***</td>
</tr>
<tr>
<td></td>
<td>(0.0444)</td>
<td>(0.0274)</td>
<td>(0.0260)</td>
<td>(0.0256)</td>
<td>(0.0418)</td>
</tr>
<tr>
<td>Observations</td>
<td>13,990</td>
<td>18,570</td>
<td>22,650</td>
<td>26,316</td>
<td>6,515</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.018</td>
<td>0.022</td>
<td>0.040</td>
<td>0.031</td>
<td>0.027</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1c father-mother-child

<table>
<thead>
<tr>
<th></th>
<th>China</th>
<th>Indonesia</th>
<th>BCS</th>
<th>HSE</th>
<th>NHANES3</th>
</tr>
</thead>
<tbody>
<tr>
<td>logbmi_father</td>
<td>0.193***</td>
<td>0.160***</td>
<td>0.200***</td>
<td>0.166***</td>
<td>0.164***</td>
</tr>
<tr>
<td></td>
<td>(0.0155)</td>
<td>(0.0108)</td>
<td>(0.0103)</td>
<td>(0.00937)</td>
<td>(0.0173)</td>
</tr>
<tr>
<td>logbmi_mother</td>
<td>0.148***</td>
<td>0.129***</td>
<td>0.176***</td>
<td>0.165***</td>
<td>0.139***</td>
</tr>
<tr>
<td></td>
<td>(0.0143)</td>
<td>(0.00891)</td>
<td>(0.00850)</td>
<td>(0.00787)</td>
<td>(0.0131)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.000***</td>
<td>2.119***</td>
<td>1.908***</td>
<td>2.127***</td>
<td>2.284***</td>
</tr>
<tr>
<td></td>
<td>(0.0594)</td>
<td>(0.0377)</td>
<td>(0.0390)</td>
<td>(0.0369)</td>
<td>(0.0624)</td>
</tr>
<tr>
<td>Observations</td>
<td>13,990</td>
<td>18,570</td>
<td>21,246</td>
<td>26,316</td>
<td>6,515</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.035</td>
<td>0.037</td>
<td>0.065</td>
<td>0.045</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
5.2 Quantile

Thus far, the estimates for IBE are at the conditional mean of child’s BMI variable. In order to explore the variation of IBE across different quantiles of child’s BMI variable (or whether the association of mother or father’s BMI variable is constant across multiple quantiles of the child’s BMI variable distribution), in this section, we estimate the quantile elasticities of BMI variable between father and child at different points in the distribution of the child’s BMI z-score.

The results are displayed in Figure 2. They suggest that the degree of BMI transmission shows a roughly increasing trend throughout child’s BMI distribution in all the samples, father-child IBE tend to be stronger at the higher level of child’s BMI variable. In other words, the effects of shared environmental and genetic factors between father and child tend to be larger for the fatter children.
Figure 2: Quantile estimates of IBE relative to OLS elasticity
5.3 Transition Matrices and Markov Chain

5.3.1 Transition Matrices

In this section, instead of using continuous variable of BMI, we use categories of BMI status as the anthropometric measure to investigate the probabilities of transition between BMI status across generations. The individuals with average BMI z-score under -1.64 are classified as underweight, -1.64~1.04 as normal weight, 1.04~1.64 as overweight, and above 1.64 as obese.
### Table 2 China

<table>
<thead>
<tr>
<th>Full sample (N= 14011)</th>
<th>Child’s BMI status</th>
<th>Mother’s distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BMI z-score</strong></td>
<td>&lt;1.64</td>
<td>-1.64-1.04</td>
</tr>
<tr>
<td><strong>Category</strong></td>
<td>Underweight</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>Mother’s BMI status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1.64</td>
<td>25.45%</td>
<td>65.45%</td>
</tr>
<tr>
<td>-1.64-1.04</td>
<td>11.3%</td>
<td>74.5%</td>
</tr>
<tr>
<td>1.04-1.64</td>
<td>6.25%</td>
<td>75%</td>
</tr>
<tr>
<td>&gt;1.64</td>
<td>6.6%</td>
<td>70.7%</td>
</tr>
<tr>
<td>Child’s distribution</td>
<td>10.68%</td>
<td>74.84%</td>
</tr>
</tbody>
</table>
### Table 3: Indonesia

<table>
<thead>
<tr>
<th>Full sample (N= 18755)</th>
<th>Child’s BMI status</th>
<th>Mother’s distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI z-score</td>
<td>&lt;-1.64</td>
<td>-1.64-1.04</td>
</tr>
<tr>
<td>Mother’s BMI status</td>
<td>Underweight</td>
<td>Normal</td>
</tr>
<tr>
<td>&lt;-1.64</td>
<td>30.1%</td>
<td>62.54%</td>
</tr>
<tr>
<td>-1.64-1.04</td>
<td>Normal</td>
<td>18.11 %</td>
</tr>
<tr>
<td>1.04-1.64</td>
<td>Overweight</td>
<td>12.47 %</td>
</tr>
<tr>
<td>&gt;1.64</td>
<td>Obese</td>
<td>9.73 %</td>
</tr>
<tr>
<td>Child’s distribution</td>
<td>16.67 %</td>
<td>71.86%</td>
</tr>
<tr>
<td>Table 4  UK (BCS 1970 cohorts)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Full sample</strong></td>
<td><strong>Child’s BMI status</strong></td>
<td><strong>Mother’s distribution</strong></td>
</tr>
<tr>
<td><strong>BMI z-score</strong></td>
<td><strong>Category</strong></td>
<td><strong>&lt;1.64</strong></td>
</tr>
<tr>
<td>&lt;1.64</td>
<td>Underweight</td>
<td>1.42 %</td>
</tr>
<tr>
<td>-1.64-1.04</td>
<td>Normal</td>
<td>1.50 %</td>
</tr>
<tr>
<td>1.04-1.64</td>
<td>Overweight</td>
<td>1.01 %</td>
</tr>
<tr>
<td>&gt;1.64</td>
<td>Obese</td>
<td>0.54 %</td>
</tr>
<tr>
<td>Child’s distribution</td>
<td></td>
<td>1.14 %</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5: UK ( HSE )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full sample</strong></td>
</tr>
<tr>
<td><strong>BMI z-score</strong></td>
</tr>
<tr>
<td>&lt;1.64</td>
</tr>
<tr>
<td>-1.64-1.04</td>
</tr>
<tr>
<td>1.04-1.64</td>
</tr>
<tr>
<td>&gt;1.64</td>
</tr>
<tr>
<td>Child’s distribution</td>
</tr>
<tr>
<td>Full sample</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td><strong>BMI z-score</strong></td>
</tr>
<tr>
<td><strong>Category</strong></td>
</tr>
<tr>
<td><strong>Mother’s BMI status</strong></td>
</tr>
<tr>
<td><strong>&lt;=-1.64</strong></td>
</tr>
<tr>
<td><strong>-1.64-1.04</strong></td>
</tr>
<tr>
<td><strong>1.04-1.64</strong></td>
</tr>
<tr>
<td><strong>&gt;1.64</strong></td>
</tr>
</tbody>
</table>
Table 2~6 describe the distribution of child’s BMI status conditional on mother’s BMI status in China, Indonesia, UK and US, respectively. This is similar to transition matrices across the discretized bivariate distribution, the classes are defined underlying the transition matrix in terms of quantile groups of the marginal distribution. The interaction terms between mother and child of different BMI status provides the matrix of intergenerational transition probabilities. For instance, in Table 2, the numbers in the first row of this matrix indicate that of the total numbers of children whose mothers were “underweight”, 25.45 percent were “underweight”, 65.45% were “normal”, 2.4% were “overweight, and 6.7% were “obese”. Similarly for the second and third row. It suggests that there is a high degree of persistence in the same BMI status across generations.

Thus, in China, the intergenerational “underweight” transmission is higher than that of other categories ; The intergenerational “underweight” transmission is even higher in Indonesia, but Indonesian mothers tend to be more overweight; The intergenerational transmission of higher BMI status ( overweight and obesity) is higher than other categories in the UK, and the BMI status tends to move up to a higher level in the offspring; These are even greater in the US.

Overall, the distribution of BMI status transition reveals a wide disparity in the strength of intergenerational transmission across different BMI status and across different countries. In underdeveloped economies, such as China and Indonesia, the intergenerational transition tends to be stronger at the lower levels of BMI status such as “underweight”; In contrast, in developed economies, such as the UK and the US, the transition tends to shift towards “normal” and higher levels of BMI status, overweight and obesity.

5.3.2 Markov Chain

As we see, Table 2~6 display the matrices of BMI status transition probabilities across generations in different countries. Based on these matrices, we build up a Markov chain to predict the future changes in the distribution of BMI status.

Using the current child’s BMI distribution as the initial distribution and based on these probabilities transition matrices, the BMI distribution in three generations is obtained by multiplying them in the following way:

Initial distribution : a= 0.1070 0.7480 0.0710 0.0810
Transition matrix: 
\[
\begin{array}{cccc}
0.2550 & 0.6550 & 0.0240 & 0.0670 \\
0.1130 & 0.7450 & 0.0660 & 0.0760 \\
0.0630 & 0.7500 & 0.0940 & 0.0930 \\
0.0660 & 0.7070 & 0.1090 & 0.1180 \\
\end{array}
\]

\[a \ast b \ast b \ast b = 0.1243 \quad 0.7365 \quad 0.0666 \quad 0.0799\]

When we multiply the initial distribution by multiple powers of transition matrices, we find that “the rows even of the third power of the basic matrix of transition probabilities are already getting close to one another”. When we raise the matrix to higher powers, we find that the Markov chain process of this transition matrix inevitably moves towards an equilibrium situation, where the constant probability vector would be

\[a \ast b \ast b \ast b \ast b = 0.1244 \quad 0.7367 \quad 0.0666 \quad 0.0799 \quad \text{(the 5th higher power)}\]

This vector shows what the proportions in different BMI status tend to be---and to remain---when the process has reached equilibrium. We find this vector (the future distribution) does not go further from the current distribution (the child’s BMI distribution in the CHNS sample), this indicates that the transition process may have already been following these matrices for a while, so that the current distribution has already reached somewhere near the equilibrium. In other words, according to the Markov Chain, the BMI distribution in China may remain as the current state in five generations if there is not a “shock” (by which I mean a dramatic environmental change, such as the food innovation) which might fundamentally change the transition matrices, or if there are no unobserved status-specific effects.

Likewise, we predict the future distribution in five generations for other countries. The results are displayed in Table 7, along with the initial distribution (the child’s BMI distribution) as a comparison. We can see that as in the CHNS sample, in other countries the distribution in five generations is also quite close to the initial distribution. However, the process of iteration varies across countries. The distribution in China and UK converges after three generations, whereas those in Indonesia and the US does not show a clear tendency of convergence.

---

5 The reason why we choose five iteration is that through the iteration of matrix, it seems that five is the optimal number of iterations through which we can see whether the resultant vector move towards convergence or not.
Overall, based on the Markov Chain, the future distribution does not vary much from the initial distribution. If this proximation of future distribution to the current distribution can be viewed as equilibrium, this may indicate that the current distribution has reached equilibrium in almost all the countries. However, whether this process converges and how long it takes to converge varies across countries. And more importantly, our results are based on the strong assumptions of Markov Chain.

Table 7: The initial distribution (child’s BMI distribution) and final distribution after five generations

<table>
<thead>
<tr>
<th>Country</th>
<th>Underweight</th>
<th>Normal</th>
<th>Overweight</th>
<th>Obese</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>China</strong></td>
<td>Initial</td>
<td>0.107</td>
<td>0.748</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.124</td>
<td>0.737</td>
<td>0.067</td>
</tr>
<tr>
<td><strong>Indonesia</strong></td>
<td>Initial</td>
<td>0.167</td>
<td>0.719</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.196</td>
<td>0.693</td>
<td>0.039</td>
</tr>
<tr>
<td><strong>UK (BCS cohorts)</strong></td>
<td>Initial</td>
<td>0.036</td>
<td>0.773</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.035</td>
<td>0.706</td>
<td>0.095</td>
</tr>
<tr>
<td><strong>Spain</strong></td>
<td>Initial</td>
<td>0.040</td>
<td>0.801</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.038</td>
<td>0.803</td>
<td>0.105</td>
</tr>
<tr>
<td><strong>US</strong></td>
<td>Initial</td>
<td>0.011</td>
<td>0.407</td>
<td>0.179</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.010</td>
<td>0.396</td>
<td>0.177</td>
</tr>
<tr>
<td><strong>HSE (pooled)</strong></td>
<td>Initial</td>
<td>0.017</td>
<td>0.550</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.019</td>
<td>0.557</td>
<td>0.148</td>
</tr>
<tr>
<td><strong>HSE (1995)</strong></td>
<td>Initial</td>
<td>0.015</td>
<td>0.550</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.014</td>
<td>0.548</td>
<td>0.140</td>
</tr>
<tr>
<td><strong>HSE (2010)</strong></td>
<td>Initial</td>
<td>0.020</td>
<td>0.530</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>Final</td>
<td>0.040</td>
<td>0.529</td>
<td>0.141</td>
</tr>
</tbody>
</table>

Table 8: HSE 1995 after one generation, compared with HSE (2010)

<table>
<thead>
<tr>
<th>Country</th>
<th>Underweight</th>
<th>Normal</th>
<th>Overweight</th>
<th>Obese</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HSE (1995)</strong></td>
<td>Initial</td>
<td>0.015</td>
<td>0.550</td>
<td>0.140</td>
</tr>
<tr>
<td></td>
<td>After one generation</td>
<td>0.014</td>
<td>0.550</td>
<td>0.140</td>
</tr>
<tr>
<td><strong>Compare with HSE (2010)</strong></td>
<td>Initial</td>
<td>0.020</td>
<td>0.530</td>
<td>0.140</td>
</tr>
</tbody>
</table>
6. Conclusions and Policy Implications

This paper has examined the intergenerational transmission of BMI or adiposity across generations in four countries across the world. Using the BMI as a measure of adiposity, we find that the intergenerational transmission of adiposity is relatively high and very comparable across time and countries – even if these countries are at a very different stage of economic development. This suggests that the intergenerational transmission mechanism which naturally occurs in biology via the transmission of both parental genetic inheritance and also the shared environment of the family when the child is growing up determine a significant fraction of the child’s likely BMI makeup as an adult. At the mean of the distribution we find that the father and mother each separately account for around 20% of the child’s BMI. Since this effect is linear and additively separable for these two parents then we have found that the joint effect of the family and its associated genetic makeup accounts for around 40% of the child’s likely BMI.

Our second key finding is that this intergenerational transmission mechanism is very different across the distribution of children’s BMI. Most specifically it is up to double for the fattest children what it is for the thinnest children. This has enormous consequences for the health of the world’s children. Specifically we find that over 30% of the fattest child’s BMI is determined by the mother and 30% by the father. Hence jointly they account for up to 60% of the child’s likely BMI. In contrast this corresponding fraction is only around 30% for the thinnest child.

To sum up, our evidence from different countries’ data suggests that there is a strong consistency in the IBE estimates across countries and ethnicities, this consistency is different from what the previous studies find on IIE and IEE. The literature on the transmission of intergenerational elasticity has found that there is a substantial disparity in the IIE and IEE estimates across different countries and different datasets. Ranging from as little as .1 to as much as .6 when they consider only the relationship of income of the child with the income of a single parent.

This first important implication of our research is that it puts the emphasis firmly on the family in terms of understand the huge fraction of adiposity determination. Specifically we need to look no further than the simply biological process of genetic inheritance from parents to child and what happens to the child when they are very young to explain a huge fraction of
what they become as fat or thin adults. We have no way (with the data available to us) of splitting up the IBE into that which is due to genetic inheritance and that which is due to the family environment – but what we do know is that jointly these two influences determine a sizeable faction of what can happen to children. One way of thinking about this process is to suggest that – in the extreme – the thinnest child in the data – who is clearly not fat at all still inherits 35% of their BMI from their parents – so that this is the lowest bound on how much may be due to the biological process of inheritance. Some fraction of the difference between their inheritance and that of the fat child with a .6 elasticity may still be due to biology but it seems likely that this could be more to do with what goes on inside the family – namely how much exercise is taken, what the family diet is like; whether they use a car for transport, how much TV is watched and generally how active they are.

The corollary of our findings is that logically – certainly for fat children much of the damage is done a the beginning of their lives. Most children are doomed to have the BMI which is pre-programmed into them by the genes they inherit and the lifestyle their family lead when they are a child. This means that the is a strictly limited amount any public intervention can do to promote health later in life. Much of the damage will have already been done.
Appendix (Data description)

CHNS data

The Chinese data here uses the longitudinal data from eight waves (1989, 1991, 1993, 1997, 2000, 2004, 2006, and 2009) of the China Health and Nutrition Survey (CHNS). Based on the definition of response rates that those who participated in previous survey rounds remaining in the current survey (Popkin, 2010), the response rates of this data were 88% at individual level and 90% at household level. This data contains detailed information on health outcomes, demographic, anthropometric measures of all members of the sampled households, including height and weight. It is noteworthy that these anthropometric measures are medically measured rather than self-reported which are mostly used in the literature. In addition, it includes information on economic and non-economic indicators such as education, household income and labor market outcomes.

Our sample is restricted to children under 18 years old with information (especially anthropometric information) on both the biological father and mother. We choose 18 as the threshold since age 18 is used to distinguish between adult and child in the CHNS physical exam dataset where the anthropometric information is included. Additionally, children within this age range normally live with their parents and rely on their parents for nutritional intake and health care. As a result, this sample includes 14,082 person-wave observations made up by 6,045 children with 3,975 fathers and 3,974 mothers. In other words, our sample includes 6,045 sets of father, mother and children.

Indonesian Family Life Survey (IFLS)

The Indonesian Family Life Survey (IFLS) is an on-going longitudinal survey data which started in 1993. The sample used here is drawn from 1993, 2000 and 2007 waves of the survey, it is representative of 83% of the Indonesian population and contains over 30,000 individuals living in 13 of the 27 provinces in Indonesia. This survey includes a range of health measures for both parents and children. It is noteworthy that as in CHNS data, the anthropometric outcome in IFLS survey was also measured by trained nurses rather than self-reported. Additionally, the IFLS data also includes information on socioeconomic factors such as education and income. Thus, the IFLS data is similar to CHNS data in terms of the survey design and measure methods, this similarity improves the comparability of results based on these two datasets.
The sample is restricted to those aged from 0 to 14 years old in each wave and have both parents and household’s information. It is noteworthy that this is different from the CHNS data, where the children sample comprises those aged between 0 and 18 years old.

In addition, in the Indonesian Family Life Survey (IFLS), we also consider step/adopted children as the sample. The adopted or step children account for around 1% of the whole sample in each wave, for these children, the information on their parents use the step parents’ rather than biological parents’

**British 1970 cohort study**

The 1970 British Cohort Study is an ongoing follow up study of 17,200 babies born in England, Scotland, Wales and Northern Ireland between 5 and 11 April 1970 who are still living in Britain (excluding Northern Ireland). The survey was conducted when the cohorts at birth, aged 5 (in 1975), 10 (in 1980), 16 (in 1986), 26 (in 1996), 30 (in 1999-2000), 34 (in 2004-2005) and 38 (in 2008-2009). The samples at the age 5 and 10 were augmented since immigrants born in the same week were added in. In this paper we use the cohorts in the first five waves (sweeps).

At the birth, the questionnaires were completed by midwife and the supplementary information was collected from clinical records. As the cohorts got older, the approach of survey changed, parents were interviewed by the health stuff and questionnaires were completed by teachers. In terms of the anthropometric information, the height and weight were measured at the age of 10 and self-reported at the age of 26 (Shaheen, et al., 1999)

**Health Survey for England**

The Health Survey for England is designated to be nationally representative of people of different age, gender, geographic region and socio-demographic circumstances. It was started in 1991 and has been conducted annually since then. The survey combines

---

6 Comparing step/adopted children BMI with the BMI of both their biological and step parents is beyond the scope of this dissertation. Several studies using adopted children have been conducted in order to investigate whether genetic influence or environmental factors that has stronger impact on child’s obesity (Sørensen 1992; Maes 1997; and Silventoinen, 2010). The results from these studies support for the importance of genetic factors, where genetic factors had a strong effect on the variation of body mass index (BMI), whereas environment indicators only show a weak influence.

7 “The 1991 and 1992 surveys had a limited population sample of about 3,000 and 4,000 adults respectively. For 1993 to 1996 adult sample was boosted to about 16,000 to enable analysis by socio-economic characteristics and health regions. In 1995 for the first time a sample of about 4,000 children was also introduced. In the 1997 Health Survey the sample was about 7,000 children and 9,000 adults. In 1998 the sample was again about 16,000 adults and 4,000 children.”
questionnaire-based answers with physical measurements and the analysis of blood sample. Each year’s survey has a particular focus on a disease or condition or population group, but height, weight and general health are covered each year. An interview with household members is followed by a nurse visit. Thus, there are both self-reported and medically-measured height and weight in this data. In the computation of BMI z-score, we use “htval” and “wtval” in the survey which are referred to as the “valid” height and weight.

**NHANES**

The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. Four surveys of this type have been conducted since 1970:

1. 1971-75—National Health and Nutrition Examination Survey I (NHANES I);
2. 1976-80—National Health and Nutrition Examination Survey II (NHANES II);
3. 1982-84—Hispanic Health and Nutrition Examination Survey (HHANES); and
4. 1988-94—National Health and Nutrition Examination Survey (NHANES III) and
5. 1999-present--National Health and Nutrition Examination Survey (Continuous NHANES)

Note in NHANES data, there is only personal identification variable (seqn), there is no household id on the public release file, the relationship of a participant to the household reference person is not publicly released. Thus, we cannot track down the participants’ parents via father and mother’s id (as in CHNS and Indonesian data), or identify the potential parents via the household id (as in England data). In other words, there is no way to identify the parents by ID. However, in one of these surveys---NHANES III---there is a family background section in the youth file, where limited characteristics of the parents were collected, including mother and father’s height and weight. Even though their age is not available, the availability of their height and weight still provides an opportunity to compute their BMI z-score, where their age is assumed 19 years old. This assumption on age is the same as we compute parents’ z-score in the previous datasets (the adults’ age above 19 are

---

8 with the exception of dietary data, the relationship of the sample participant to the proxy is not publicly released, either.

9 Note in NHANES2, there is also mother and father’s height and weight in the youth file of on the Web of Tutorial, but most of them are the same, I suspect that they are wrong, so not use them
assumed 19 when they run through the WHO macro). Therefore, NHANES III allows us the

NHANES III, conducted between 1988 and 1994, included about 40,000 people selected
from households in 81 counties across the United States. In NHANES III, black Americans
and Mexican Americans were selected in large proportions, each of these groups comprised
separately 30 percent of the sample. It was the first survey to include infants as young as 2
months of age and to include adults with no upper age limit. Our sample is obtained by
merging the youth data which includes child’s age and parents’ height weight with
examination data which includes child’s final (medically measured) height weight. Any BMI
zscores outside the range of -6 and 6 are viewed as implausible and are dropped from the
sample (Li et al., 2009). Our final sample includes 6,582 pairs of father, mother and child.
References


Corak, M. and A. Heisz (1999), The Intergenerational Earnings and Income Mobility of


Thompson, O. The Intergenerational Transmission of Health Status: Estimates and Mechanisms.


